

Wed May 29 16:25:20 2002

us-09-676-436-3.szlml00.rn1

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 10:15:51; Search time 94.14 Seconds  
(without alignments)  
13020.089 Million cell updates/sec

Title: US-09-676-436-3  
Perfect score: 4990  
Sequence: 1 ctgaagactcttcgacatga.....ggtattattagaagctc 4990

Scoring table: IDENTITY\_MDC  
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues  
Total number of hits satisfying chosen parameters: 613726

Minimum DB seq length: 0  
Maximum DB seq length: 100  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTC.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/PTC.COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	26	0.5	100	US-08-400-256-4
2	26	0.5	100	US-08-975-365-4
3	25.2	0.5	99	US-08-427-097-5
4	25.2	0.5	99	US-08-878-957-5
5	24.6	0.5	71	US-08-458-423A-78
6	24.6	0.5	71	US-08-458-424B-78
7	24.6	0.5	71	US-08-973-124-78
8	24.6	0.5	71	US-08-973-124-78
9	24.6	0.5	71	US-08-973-124-78
10	24.4	0.5	100	US-08-400-256-9
11	24.4	0.5	100	US-08-400-256-9
12	24.4	0.5	100	US-08-400-256-9
13	24.4	0.5	100	US-08-400-256-9
14	24.4	0.5	100	US-08-400-256-9
15	23.8	0.5	97	US-08-182-175A-54
16	23.6	0.5	97	US-08-182-175A-54
17	23.6	0.5	97	US-08-182-175A-54
18	23.4	0.5	97	US-08-182-175A-54
19	23.4	0.5	97	US-08-182-175A-54
20	22.8	0.5	96	US-08-484-322-5
21	22.8	0.5	96	US-08-484-322-5
22	22.6	0.5	97	US-07-679-052A-10
23	22.4	0.4	90	US-08-123-702-21
24	22.4	0.4	95	US-08-332-766A-41
25	22.4	0.4	97	US-08-182-175A-48
26	22.4	0.4	97	US-08-474-633A-57
27	22.4	0.4	97	PCT-US92-06412-48

28	22.2	0.4	71	US-08-465-591A-75	Sequence 75, Appl
29	22.2	0.4	71	US-08-465-594A-75	Sequence 75, Appl
30	22.2	0.4	71	US-08-973-124-260	Sequence 260, App
31	22.2	0.4	71	PCT-US96-08014-260	Sequence 99, Appl
32	22.2	0.4	85	US-08-488-402A-99	Sequence 99, Appl
33	22.2	0.4	85	US-08-484-552A-99	Sequence 99, Appl
34	22.2	0.4	85	PCT-US96-09472-99	Sequence 3, Appl
35	22.2	0.4	87	US-08-110-286A-3	Sequence 166, App
36	22.2	0.4	90	US-08-461-697-166	Sequence 166, App
37	21.8	0.4	54	US-08-758-306-514	Sequence 114, App
38	21.8	0.4	79	US-08-933-983-55	Sequence 57, Appl
39	21.8	0.4	79	US-08-933-983-55	Sequence 57, Appl
40	21.6	0.4	81	US-08-445-463B-10	Sequence 10, Appl
41	21.6	0.4	65	US-08-445-464C-10	Sequence 10, Appl
42	21.6	0.4	65	PCT-US94-03437-10	Sequence 23, Appl
43	21.6	0.4	85	US-08-602-093-23	Sequence 113, App
44	21.6	0.4	93	US-08-413-118-113	Sequence 113, App
45	21.6	0.4	93	US-08-473-446-113	Sequence 113, App

# ALIGNMENTS

RESULT 1  
US-08-400-256-4/c  
Sequence 4, Application US/08400256  
Patent No. 5750497  
GENERAL INFORMATION:  
APPLICANT: Havelund, Svend  
APPLICANT: Halstrom, John  
APPLICANT: Jonassen, Ib  
APPLICANT: Andersen, Asger Sloth  
APPLICANT: Markussen, Jan  
TITLE OF INVENTION: ACYLATED INSULIN  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESS: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/400/256  
APPLICATION NUMBER: US/08/400/256  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 3965.220-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9655  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-400-256-4  
Query Match 0.5%; Score 26; DB 1; Length 100;  
Best Local Similarity 55.6%; Pred. No. 6.1e+02;  
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
QY 4469 aagaatgaagcctagatagatactggaataatcttaatacactacatgatgtat 4528

Db 97 AAGCTACACGACGCTAAGGCTATCGTTGACAAATGTTACTTATCTGTTCTTGTAC 38  
QY 4529 tattacataagacgtctgtgaagcag 4558  
Db 37 CAATGGAAACTACTGTCTTGAAGCAG 8

## RESULT 2

US-08-975-365-4/C  
Sequence 4, Application US/08975365  
Patent No. 6011007  
GENERAL INFORMATION:  
APPLICANT: Havelund, Svend  
APPLICANT: Halstrom, John  
APPLICANT: Jonassen, Ib  
APPLICANT: Andersen, Asger Sloth  
APPLICANT: Markussen, Jan  
TITLE OF INVENTION: ACYLATED INSULIN  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,365  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/400,256  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 3985,220-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-867-9655  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-975-365-4

Query Match 0.5%; Score 26; DB 3; Length 100;  
Best Local Similarity 55.6%; Pred. No. 6,1e+02;  
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 4469 aagaatgaagccatgagatgactggaatcttctaactactgataa 4528  
Db 97 AAGCTGACGACGCTAAGGCTATCGTTGACAAATGTTACTTATCTGTTCTTGTAC 38  
QY 4529 tattacataagacgtctgtgaagcag 4558  
Db 37 CAATGGAAACTACTGTCTTGAAGCAG 8

RESULT 3  
US-08-427-097-5  
Sequence 5, Application US/08427097  
Patent No. 568294

GENERAL INFORMATION:  
APPLICANT: Meagher, Richard B.  
APPLICANT: Sommers, Anne O.  
TITLE OF INVENTION: Metal Resistance Sequences and  
TITLE OF INVENTION: Transgenic Plants  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/427,097  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferder, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 40-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Oligonucleotide"  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-427-097-5

Query Match 0.5%; Score 25.2; DB 1; Length 99;  
Best Local Similarity 54.3%; Pred. No. 1e+03;  
Matches 51; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3026 atgctcgaactaagtgtgaggtgaggtacaaagccaggtggcgaactcaagat 3085  
Db 2 AAGACCCAGCTATAGTGAAGCTGTACTGCTGATTCGATGGAAGCATTGAAGTGC 61  
QY 3086 ttgatttctacaagcaattgaaccttccttat 3119  
Db 62 GTGAGCATATCTCAAGCAAGCAAGTCTTATAT 95

## RESULT 4

US-08-878-957-5  
Sequence 5, Application US/08878957  
Patent No. 5965796

GENERAL INFORMATION:  
APPLICANT: Meagher, Richard B.  
APPLICANT: Summers, Anne O.  
APPLICANT: Rugh, Clayton L.  
TITLE OF INVENTION: Metal Resistance Sequences and  
TITLE OF INVENTION: Transgenic Plants  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303



FILED DATE: 11-JUNE-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/964,624  
FILING DATE: 21-OCTOBER-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/117,991  
FILING DATE: 8-SEPTEMBER-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/931,473  
FILING DATE: 17-AUGUST-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Diane H. McLearn  
REGISTRATION NUMBER: 33,960  
REFERENCE/DOCKET NUMBER: NX 34-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 71 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-458-424B-78

Query Match  
Best Local Similarity 0.5%; Score 24.6; DB 1; Length 71;  
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgtgattagctgtattcaagcagatcaccattgcgacgaagctcttc 4012  
Db 56 GAAGACGTAGTAGTAGTAGTATTAACATACCCCTGCGCCGATGCTCTC 2

RESULT 7  
US-08-973-124-78/c  
Sequence 78, Application US/08973124  
Patent No. 6207816  
GENERAL INFORMATION:  
APPLICANT: LARRY GOLD et al.  
TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE  
TITLE OF INVENTION: LIGANDS TO GROWTH  
NUMBER OF INVENTIONS: FACTORS  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 200  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,124  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/08014  
FILING DATE: 30-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/458,423  
FILING DATE: 02-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/458,424  
FILING DATE: 02-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/465,594

FILED DATE: 05-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/465,591  
FILING DATE: 05-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/479,725  
FILING DATE: 07-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/479,783  
FILING DATE: 07-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/518,693  
FILING DATE: 20-MARCH-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 71 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-973-124-78

Query Match  
Best Local Similarity 0.5%; Score 24.6; DB 4; Length 71;  
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgtgattagctgtattcaagcagatcaccattgcgacgaagctcttc 4012  
Db 56 GAAGACGTAGTAGTAGTAGTATTAACATACCCCTGCGCCGATGCTCTC 2

RESULT 8  
PCT-US96-08014-78/c  
Sequence 78, Application PC/TUS9608014  
GENERAL INFORMATION:  
APPLICANT: LARRY GOLD; NEBOJSA JANIC; STEVEN RINGQUIST; NIKOS  
TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE  
TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH  
TITLE OF INVENTION: FACTOR (TGF), PLATELET-DERIVED  
TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN  
NUMBER OF INVENTIONS: 304  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 200  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/08014  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/458,423  
FILING DATE: 02-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/458,424  
FILING DATE: 02-JUNE-1995

? SEQUENCE 9, APPLICATION US/08400256  
 ? PATENT NO. 5750497  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Havelund, Svend  
 ? APPLICANT: Halstrom, John  
 ? APPLICANT: Jonassen, Ib  
 ? APPLICANT: Andersen, Asger Sloth  
 ? APPLICANT: Markussen, Jan  
 ? TITLE OF INVENTION: ACYLATED INSULIN  
 ? NUMBER OF SEQUENCES: 49  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESSEE: No. 57504970 No. 5750497disk of No. 5750497th America, Inc  
 ? STREET: 405 Lexington Avenue, 64th Floor  
 ? CITY: New York  
 ? STATE: New York  
 ? COUNTRY: United States of America  
 ? ZIP: 10174-6401  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: Patent In Release #1.0, Version #1.25  
 ? SOFTWARE: Application DATA:  
 ? CURRENT APPLICATION NUMBER: US/08400,256  
 ? FILING DATE: 03-MAR-1995  
 ? CLASSIFICATION: 514  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Lambiris, Elias J.  
 ? REGISTRATION NUMBER: 33,728  
 ? REFERENCE/DOCKET NUMBER: 3985,220-US  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: 212-867-0123  
 ? TELEFAX: 212-878-9655  
 ? TELEFAX FOR SEO ID NO: 9:  
 ? INFORMATION CHARACTERISTICS:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 100 base pairs  
 ? TYPE: nucleic acid  
 ? STRANDEDNESS: single  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: DNA  
 ? US-08-400-256-9

Query Match 0.5%; Score 24.4; DB 1; Length 100;  
 Best Local Similarity 54.4%; Pred. No. 1.7e+03;  
 Matches 49; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 4469 aagaatgaagcctagatagatagactggaataatctcttaacactactgatatga 4528  
 DB 97 AAGCTGACGACGACCTAAGGATATCGTTGAACAATGTGACTCTATCTGTCTTTGAC 38  
 OY 4529 tattacataaagactgctgtagaagcag 4558  
 DB 37 CAATTGGAATACTACTGTCTGTAGACGAG 8

## RESULT 11

US-08-975-365-9/c  
 ; Sequence 9, Application US/08975365  
 ; Patent No. 6011007  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Havelund, Svend  
 ; APPLICANT: Halstrom, John  
 ; APPLICANT: Jonassen, Ib  
 ; APPLICANT: Andersen, Asger Sloth  
 ; APPLICANT: Markussen, Jan  
 ; TITLE OF INVENTION: ACYLATED INSULIN  
 ; NUMBER OF SEQUENCES: 49  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.  
 ; STREET: 405 Lexington Avenue, 64th Floor  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 10174-6401  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/975,365  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/400,256  
 ; FILING DATE: 03-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lambiris, Elias J.  
 ; REGISTRATION NUMBER: 33,728  
 ; REFERENCE/DOCKET NUMBER: 3985,220-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-867-0123  
 ; TELEFAX: 212-878-9655  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 100 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA  
 ; US-08-975-365-9

Query Match 0.5%; Score 24.4; DB 3; Length 100;  
 Best Local Similarity 54.4%; Pred. No. 1.7e+03;  
 Matches 49; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 4469 aagaatgaagcctagatagatagactggaataatctcttaacactactgatatga 4528  
 DB 97 AAGCTGACGACGACCTAAGGATATCGTTGAACAATGTGACTCTATCTGTCTTTGAC 38  
 OY 4529 tattacataaagactgctgtagaagcag 4558  
 DB 37 CAATTGGAATACTACTGTCTGTAGACGAG 8

## RESULT 12

US-08-182-175A-54  
 ; Sequence 54, Application US/08182175A  
 ; Patent No. 5559223  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saverio Carl Falco  
 ; APPLICANT: Sharon J. Keeler  
 ; APPLICANT: Janet A. Rice  
 ; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containin  
 ; NUMBER OF SEQUENCES: 113  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: E.I. du Pont de Nemours and Company  
 ; STREET: 1007 Market Street  
 ; CITY: Wilmington  
 ; STATE: Delaware  
 ; COUNTRY: USA  
 ; ZIP: 19898  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy Disk  
 ; OPERATING SYSTEM: Macintosh  
 ; SOFTWARE: Microsoft Word, 4.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/182,175A  
 ; FILING DATE:  
 ; CLASSIFICATION: 800  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/743,006  
 ; FILING DATE: 9 August 1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Linda Axamethy Floyed  
 ; REGISTRATION NUMBER: 33,692  
 ; REFERENCE/DOCKET NUMBER: BB-1031  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (302) 982-4929  
 ; TELEFAX: (302) 892-7949  
 ; TELEX: 835420  
 ; INFORMATION FOR SEQ ID NO: 54:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 97 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; ORIGINAL SOURCE:  
 ; STRAIN: E. coli  
 ; CELL TYPE: DH5 alpha  
 ; IMMEDIATE SOURCE:  
 ; CLONE: 92-2  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 2..88  
 ; OTHER INFORMATION: /function="synthetic storage protein  
 ; OTHER INFORMATION: /product="protein"  
 ; OTHER INFORMATION: /gene="ssp"  
 ; OTHER INFORMATION: /standard\_name="5.11.11.5"  
 ; US-08-182-175A-54

Query Match 0.5%; Score 24; DB 1; Length 97;  
 Best Local Similarity 58.3%; Pred. No. 2.2e+03;  
 Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 1120 gagatagagggtgacacagaagaaggaataaagaggtggaagtagtagatgag 1179  
 DB 10 GAGATGATGAGGCGATGAGAGAGAGATGAGATGAGAGAGATGAGATGAGATGAG 69  
 OY 1180 agtgaagaaga 1191  
 DB 70 AGAGAGATGAA 81

```

RESULT 14
PCT-US92-06412-54
; Sequence 54, Application PC/TUS9206412
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containin
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh System, 6.0
; OPERATING SYSTEM: Macintosh Word, 4.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06412
; FILING DATE: 19920807
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamechy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: E. coli
; CELL TYPE: DH5 alpha
; IMMEDIATE SOURCE:
; CLONE: 92-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..98
; OTHER INFORMATION: /function= "synthetic storage protein
; OTHER INFORMATION: /product= "protein"
; OTHER INFORMATION: /gene= "ssp"
; OTHER INFORMATION: /standard_name= "5.11.11.5"
; OTHER INFORMATION:
;
PCT-US92-06412-54

Query Match 0 5%; Score 24; DB 5; Length 97;
Best Local Similarity 58.3%; Pred. No. 2.2e+03;
Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 1120 gagatctaggttgattgacacagaagaagaaatlaaaggagtctgaagaagtactgcgatgag 1179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 GAGGATGTAAGCGCATGGAGGAGGAAGAATGAGTGCATGAGGAGAGAAAGATGAAGTGATGCA 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1180 agtgaagaagaaa 1191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 AGGAGAAATGCAA 81

```

Query Match	0.5%	Score 23.8;	DB 1;	Length 83;
Best Local Similarity	72.1%	Pred. No. 2.3e+03;		
Matches 31;	Conservative 0;	Mismatches 12;	Indels 0;	Gaps 0;
QY 1352	gttagacatctcgaatctattattatagaccacattttagacaa	1394		
Dd 17	GTACACGTTATGTGACTTATAGTTAGTCACTTCATCTGGGCCAA	59		

Search completed: May 29, 2002, 12:56:56  
Job time: 9665 sec



Wed May 29 16:25:21 2002

us-09-676-436-3.szlml100.rst

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 09:13:38 ; Search time 3518.59 Seconds  
(without alignments)  
19141.135 Million cell updates/sec

Title: US-09-676-436-3  
Perfect score: 4990  
Sequence: 1 ctgaagactctccgatagta.....ggttatttaggaagctc 4990  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 297742

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estha:\*  
2: em\_esthm:\*  
3: em\_esthm:\*  
4: em\_esthm:\*  
5: em\_esthm:\*  
6: em\_esthm:\*  
7: em\_esthm:\*  
8: em\_esthm:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pla:\*  
16: em\_gss\_vit:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	39.8	0.8	95	AA990067	AA990067 ua58f02.r
2	29.8	0.6	70	AA990364	AA990364 aa44c10.r
3	28.6	0.6	71	A1767078	A1767078 w192907.x
4	27.8	0.6	83	B1063432	B1063432 B1063432
5	27.8	0.6	83	B1063432	B1063432 B1063432
6	27.8	0.6	94	A1360096	A1360096 qy89d09.x
7	27.8	0.6	94	B1322285	B1322285 kx19a08.y
8	27.4	0.5	100	AA154655	AA154655 mq61c08.x
9	27.4	0.5	98	A1054386	A1054386 q176e03.x
10	26.8	0.5	100	C53310	C53310 C53310 yu11
11	26.6	0.5	81	BE573242	BE573242 601333139
12	26.6	0.5	82	BF101686	BF101686 601753418
13	26.4	0.5	87	AA267457	AA267457 m290a03.r
14	26.4	0.5	95	AA946634	AA946634 c038f11.s
15	26.4	0.5	95	R93104	R93104 EST000038.S
16	26.2	0.5	78	R40354	R40354 yf81d05.s1
17	26.2	0.5	93	AV551027	AV551027 AV551027

18	25.2	0.5	93	10	T62011	T62011 yb97b12.r1
19	26	0.5	78	10	BI858071	BI858071 603384887
20	26	0.5	85	12	BI127076	BI127076 G-10g24.f
21	26	0.5	88	9	AA865746	AA865746 o44h11.s
22	26	0.5	98	9	A1251245	A1251245 qv55412.x
23	25.8	0.5	78	9	A1035450	A1035450 u447b05.r
24	25.8	0.5	83	9	AA915897	AA915897 o86b08.s
25	25.8	0.5	91	9	A1340694	A1340694 lb33c05.x
26	25.6	0.5	73	9	A1894187	A1894187 mc67e06.x
27	25.6	0.5	73	9	A1142956	A1142956 c258d01.x
28	25.6	0.5	85	9	A1086378	A1086378 o244c01.x
29	25.6	0.5	86	9	AV532477	AV532477 AV532477
30	25.4	0.5	79	9	AA149995	AA149995 x032c06.x
31	25.4	0.5	89	12	A2961698	AA149995 x032c06.x
32	25.4	0.5	91	12	D19910	AA2610198 2M0230B17
33	25.4	0.5	92	12	AF149669	AA2610198 2M0230B17
34	25.4	0.5	93	9	AA669624	D1991069 AF149669
35	25.4	0.5	94	9	AA662095	AA669624 ac20a02.s
36	25.4	0.5	94	9	AA662095	AA662095 s
37	25.4	0.5	95	9	AA576542	AA662095 s
38	25.4	0.5	99	10	BC632479	AA576542 rmb6c06.s
39	25.2	0.5	99	10	BE054677	AA576542 rmb6c06.s
40	25.2	0.5	64	12	AE66720	BE054677 7169e05.y
41	25.2	0.5	81	9	AA451873	BE054677 7169e05.y
42	25.2	0.5	88	9	AA451873	AE66720 SMOVAE008
43	25.2	0.5	88	9	AA451873	AE66720 SMOVAE008
44	25.2	0.5	94	10	AA451873	AE66720 SMOVAE008
45	25.2	0.5	97	12	TA330B01P	AE66720 SMOVAE008

# ALIGNMENTS

RESULT 1  
AA990067 95 bp mRNA linear EST 02-JUN-1998  
ua58f02.r1 Scores: rhyms\_2NBMt Mus musculus CDNA clone  
IMAGE:1361691 5', mRNA sequence.  
AA990067.1 GI:3175431  
EST.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 95)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, U., Moore, B.,  
Scheltemberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,  
Theisinger, B., Wylie, J., Lennon, G., Soares, B., Willson, R. and  
Waterston, R.  
The WashU-HMNI Mouse EST Project  
Unpublished (1996)  
Contact: Maria M/Mouse EST Project  
WashU-HMNI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through INLM; contact the  
IMAGE Consortium (info@image.jlnl.gov) for further information.  
MGI:699731  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 82.  
Location/Qualifiers  
1..95  
/organism="Mus musculus"  
/strain="C57BL/6J1090"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1361691"  
/clone\_id="Soares-Chymus\_2NBMt"  
/sex="male"





' Gibbons, R., Ritter, E., Bennett, J., Franklin, C., Tsagaris, S.,  
 ' Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K.,  
 ' M. Allen, M., Pearson, B., Swailier, T., Harey, N.,  
 ' Shint, T., Jackson, Y., Cardenas, M., McCann, R.,  
 ' Wilson, R., Jackson, Y., Cardenas, M., McCann, R.,  
 ' The Washington Univ. Nematode EST Project, 1999  
 ' Unpublished (1999)  
 ' Contact: MOC (1999)

FEATURES  
SOURCE

Fax: 314 286 1800  
 Tel.: 314 286 1810  
 Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
 The library was constructed by Brandi Chiapelli and Dr. James McCarter ([jchapel@watson.wustl.edu](mailto:jchapel@watson.wustl.edu) & [jmcarter@watson.wustl.edu](mailto:jmcarter@watson.wustl.edu)) at Washington University, St. Louis. DNA Sequencing by: Washington University Genome Sequencing Center St. Louis.  
 Putative full length read  
 The vector to vector length is 95  
 Seq primer: -40RP from Gibco  
 Location/Qualifiers

BASE COUNT  
ORIGIN

42 a (walwick.grant@agresearch.co.nz),<sup>h</sup> research, New Zealand  
9 c  
12 g  
31 t

```
Query Match
Best Local
Matches 4
```

```

0.0%; Score 27.8; DB 10; Length 94;
62.0%; Pred. No. 3.1e+04;
0; Mismatches 27; Indels 0; Gaps 0;

```

D8  
TTTATCAAGTT 78

RESULT	7
AA154655/c	
LOCUS	
DEFINITION	AA
	mq
	5'
ACCESSION	PR
VERSION	AA
KEYWORDS	AA1
SOURCE	EST
ORGANISM	hou
	Mus

*Scares*, thymus, 100 bp mRNA linear EST  
5' similar to SW:KCC4.MOUSE P08R14 CALCIUM/  
PROTEIN KINASE TYPE IV CATALYTIC CHAIN ;  
AA154655.1 GI:1726483 mRNA sequence.  
EST.

REFERENCE  
AUTHORS

1 (bases 1 to 100)  
Marra, M., Luchetta; Rodentia; Sciurognathi; Muridae; Eulestomini;  
Miller, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Celsel, S., Knab, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

TITLE  
JOURNAL  
COMMENT

Waterston, R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
Contact: Maria M/Mouse.com

FEATURES  
source

This clone is available royalty-free through LMU ; contact the IMAG Consortium (info@image.lmu.gov) for further information. MG13357862

Trace considered overall poor quality  
Possible reversed clone: similarity  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 1.

BASE COUNT  
ORIGIN

24 a Soares and M. Fatima Bonaldo.  
11 c  
20 g 45 t

Query Match  
Best Local

57.58;	score 2/.8;	DB 9;	Length 100;
Pred. No. 3.2e+04;			

	Conservative	Mismatches	Indels	Gaps
Qy 3977	atcaaaagcagatcattcgcattcaaaagctccatcgaatgacatgcatgctccacggt	0	37	0
Db 87	ATCACAAGCAGACATTAGAGGGCATTCATATTCTTACGACAAATMAATATCATCCACGAG	11	11	11
Qy 4037	acattaaagtgccacatcttccctta	4	6	3
Db 27	ATCTAAAGCTAGCAATATTCTCTATA	1	11	1

RESULT	8
T85845/c	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	

T85845 91 bp mRNA linear EST 17-MAR-1995  
Yd61010.r1 Soares fetal liver spleen INRIS Homo sapiens cDNA clone  
IMAGE:117723 similar to gb:X56468\_rna1.14-3-3 PROTEIN THETA  
(HUMAN)., mRNA sequence.  
T85845 EST.  
T85845.1 GI:714197  
human.

REFERENCE  
AUTHORS

Reproductive behavior of the following groups of mammals: Eutheria; Primates; Carnivora; Eutelestomida 1 (Bases 1 to 91)  
 Biller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
 Riffin, L., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
 R., Williamson, A., Wohlmann, P. and Wilson, T.



BASE COUNT	54 a	7 c	18 g	21 t
ORIGIN				

Query Match	0.5%;	Score 26.8;	DB 10;	Length 100;
Best Local Similarity	55.3%;	Pred. No. 5.9e+04;		
Matches 52;	Conservative	0;	Mismatches 42;	Indels 0.

QY 131 cttattggaagaaaaatgaatgcacaaatcagccctccacataaagacnctggaaaaa 190  
 Db 3 cttttttgaaagacgaacaaatataagaaaataaaattaaaaaaatttagcaaaatgcgaanng 62  
 QY 191 cagtggaagatttggaagaatacagctataaagca 224  
 Db 63 ttgcgattcagctcggaagaaaacattgacttgaa 96

RESULT	11
LOCUS	BE573242
DEFINITION	81 bp mRNA linear EST 15-AUG-2000
ACCESSION	601333139F2 NCI-CGAP_Mam6 Mus musculus cDNA clone IMAGE:3710769 5'
VERSION	BE573242
KEYWORDS	BE573242.1 GI:9816962
SOURCE	EST.
ORGANISM	house mouse. Mus musculus

REFERENCE	Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus
AUTHORS	1 (bases 1 to 81)
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Robert Strassberg, Ph.D.

```

: name: LHAM8933 row: m column: 10
High quality sequence stop: 81.
Location/Qualifiers
1..81
FEATURES
source
```

BASE COUNT	25 a	9 c	13 g	34 t
ORIGIN	Providing samples: Jeffrey Green, M.D., NIH			

Query Match	0.58;	Score 26.6;	DB 10;	Length 81;
Best Local Similarity	63.1%;	Pred. No. 6.1e+04;		
Matches	41; Conservative	0; Mismatches	24;	

OY 4922 aaaggtgcttcacbaaatlaagaanaaggctcttttccataaaagtattatcta 4981  
| | | | | | | | | | | | | | | | | |  
db 6 ATAAAGGCTGGTACTAGAGTTAATGCACAAATGTTCATGCGCATTTCTTTTATAATG 65  
OY 4982 ggaag 4986

Db 66 TTA 70

RESULT	12
BFI01686	
LOCUS	BFI01686
DEFINITION	601753418r1 NCI-GAP_Mam1 Mus musculus cDNA clone IMAGE:3980885 5'
ACCESSION	MRI01686

VERSION	BE101686.1	GI:10884212
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 82)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: [cgabods-r@mail.nih.gov](mailto:cgabods-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: L14M9177 row: d column: 06  
High quality sequence step: 82.  
Location/Qualifiers  
1: 82  
source

BASE COUNT	ORIGIN
24 a	9 c 13 g 36 t

```

Query Match      0.58;   Score 26.6;   DB 10;   Length 82;
Best Local Similarity 63.1%;   Pred No. 6.1e+04;
Matches 41;   Conservative 0;   Mismatches 24;   Indels 0;   Gaps 0;

```

*SJ* 4972 ddaaggcgtattactggaatttagaaaaggtcttttccataataygttatllta 4987  
| | | | | | | | | | | | | | | | | | |  
*Dd* 5 ATAAGCTCCTTACTAGAGTTTAATGCAAACTTCATGCCTCATTTCTTTTTATCATG 64  
| | | | | | | | | | | | | | | | | | |  
*OY* 4982 ggaa 4986  
| | |  
*Dd* 65 TTAAA 69

RESULT 13  
AA267457

DEFINITION	m29a03.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
ACCESSION	AA267457
	87 bp mRNA linear EST 21-MAR-1997

VERSION	AA26/457.1	GI:1904190
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



Wed May 29 16:25:21 2002

us-09-676-436-3.sz1m100.rst

Email: M-Saber@RCU.EUN.BG  
Seq primer: SK

FEATURES  
source

location/Qualifiers  
1..95

/organism="Schistosoma mansoni"

/strain="Egyptian"

/db\_xref="taxon:6183"

/clone="SMTBADAMS0038SK"

/clone\_lib="S. mansoni cDNA"

/lab\_host="E. coli XL Blue1"

/note="vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2: XhoI; mRNA was constructed and cloned simultaneously using vector priming with the pBluescript II SK+ vector. cDNA was directionally synthesized from the EcoRI site in the vector to the XhoI site."

BASE COUNT 31 a 15 c 10 g 39 t  
ORIGIN

Query Match

Best Local Similarity 59.2%; Score 26.4; DB 10; Length 95;  
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

OY 641 actctcttatacaagccgtaagccatccagatatataatgaatcctactt 700  
||| || |||| | || |||| ||| |||| |||| |||| |||| ||||  
Db 5 ACTATGATATATATAACACTACTTCAGCAATCCCTATCTATAAATGTATCTTATTT 64  
OY 701 tcaagtcgactatgg 716  
| |||| | ||||  
Db 65 TGAATTTTACCCTATG 80

Search completed: May 29, 2002, 11:17:02  
Job time: 7404 sec



Wed May 29 16:25:19 2002

us-09-676-436-3.sz1ml100.rng

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 10:21:32 ; Search time 487.5 Seconds  
(without alignments)  
17574.160 Million cell updates/sec

Title: US-09-676-436-3  
Perfect score: 4990  
Sequence: 1 ctgaagactctccgatga.....gtttatttagaagctc 4990

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues  
Total number of hits satisfying chosen parameters: 2046006

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	27.2	0.5	78	18	AAT50934	Mouse p53-recognt
2	26.2	0.5	92	21	AAQ19445	Human secreted pro
3	25.2	0.5	100	16	AAQ86400	Human insulin deri
4	25.2	0.5	60	22	AAH48689	Human G-protein su
5	25.2	0.5	70	21	AAH71594	Human brain natu
6	25.2	0.5	89	16	AAT22086	Human gene signat
7	25.2	0.5	99	18	AAT92244	Mercuric ion reduc
8	25.2	0.5	99	20	AAZ28594	Primer 307-3395 fo
9	25	0.3	54	21	AAH73946	GFP Leu(CITG) 5 forw

10	25	0.5	91	22	ABA69691	Human foetal liver
11	24.8	0.5	77	21	AAQ12392	Human secreted pro
12	24.8	0.5	92	21	AAQ06930	Reverse PCR primer
13	24.8	0.5	94	21	AAQ11499	Human secreted pro
14	24.8	0.5	94	21	AAQ15541	Human secreted pro
15	24.6	0.5	71	18	AAH65208	Transforming growt
16	24.6	0.5	72	21	AAH52344	CDNA encoding dros
17	24.6	0.5	98	20	AAH23468	Human neutrophil c
18	24.4	0.5	82	20	AAH245128	Human foetal liver
19	24.4	0.5	99	22	AAH68615	Probe #14054 for g
20	24.4	0.5	99	22	AAH68615	Human insulin deri
21	24.4	0.5	100	16	AAQ86411	Human foetal liver
22	24.2	0.5	80	22	AAH72501	Human brain expres
23	24.2	0.5	80	22	AAH71381	Human bone marrow
24	24.2	0.5	80	22	AAH73388	Human bone marrow
25	24.2	0.5	80	22	AAH73326	Probe #21912 used
26	24.2	0.5	84	21	AAH92397	Rhymosin alpha-1 n
27	24.2	0.5	94	21	AAH51645	Spd1 exon sequence
28	24.2	0.5	94	21	AAH51658	Spd1 exon sequence
29	24.2	0.5	94	21	AAH51692	Human foetal liver
30	24.2	0.5	95	22	AAH71915	Human brain expres
31	24.2	0.5	95	22	AAH20302	Human bone marrow
32	24.2	0.5	95	22	AAH46381	Probe #20956 used
33	24.2	0.5	95	22	AAH52270	Human foetal liver
34	24.2	0.5	99	22	AAH75243	Human bone marrow
35	24.2	0.5	99	22	AAH39902	Human brain expres
36	24.2	0.5	99	22	AAH23782	Human bone marrow
37	24.2	0.5	99	22	AAH49875	Human brain expres
38	24.2	0.5	99	22	AAH26980	Probe #16913 for g
39	24.2	0.5	99	22	AAH5798	Probe #24484 used
40	24	0.5	66	22	AAH14229	Synthetic transcri
41	24	0.5	78	24	AAH63113	Cell death protecl
42	24	0.5	97	14	AAQ37278	Clone 92-2 used in
43	24	0.5	97	16	AAQ94997	SP 5.11.11.5 clon
44	24	0.5	97	19	AAH5835	Synthetic storage
45	24	0.5	97	20	AAH95532	Synthetic lysine-r

#### ALIGNMENTS

RESULT 1  
AAT50934 standard; CDNA: 78 BP.  
ID AAT50934:  
XX  
AC AAT50934:  
XX  
XX 07-APR-1997 (first entry)  
XX  
XX Mouse p53-recognition clone 1 5' sequence.  
XX  
XX p53 binding protein; Mdmx; tumour suppressor; cancer; ss.  
XX  
XX Mus sp.  
XX  
XX W09641875-A1.  
XX  
XX 27-DEC-1996.  
XX  
XX 13-JUN-1996; 96W0-NL00239.  
XX  
XX 13-JUN-1995; 95EP-0201565.  
XX  
XX (INTR-) INTRIGENE BV.  
XX (OTLE-) RIKUSUNIV LEIDEN.  
XX  
XX Jochimsen A, Shwarts A, Van Der Eb AJ;  
XX  
XX WPI; 1997-065462/06.  
XX  
XX Nucleic acid encoding p53-binding protein - which has homology to  
XX  
XX mouse Mdm2, for use in cancer research

PS Disclosure; Fig 2a; 30pp; English.

CC The sequences of the 5' region (AAT50934) and 3' region (AAT50935)  
 CC were derived for cDNA clone 1, isolated from a mouse 16-day embryo.  
 CC cDNA library using tumour suppressor p53 as probe. The 5' and 3'  
 CC regions (AAT50936-37) of another clone were also derived. These  
 CC sequences showed no significant homology to any known sequence  
 CC submitted to several databases. A third clone (see also AAT50933)  
 CC coded for a p53 binding protein designated Minx (AAT10206).  
 XX Sequence 78 BP; 33 A; 9 C; 22 G; 10 T; 4 other;  
 SQ

Query Match  
 Best Local Similarity 0.5%; Score 27.2; DB 18; Length 78;  
 Matches 43; Conservative 2; Mismatches 29; Indels 0; Gaps 0;

QY 1122 gtaataaggggtgacacagaagaattgaaggagttggaagtagtgcgagatgag 1181  
 Db 1 gtaataaggggtgacacagaagaattgaaggagttggaagtagtgcgagatgag 1181  
 QY 1182 tgaagaagaacaa 1195  
 Db 61 tgvnhaagacaaa 74

# RESULT 2

AAC19445/c  
 ID AAC19445 standard; cDNA; 92 BP.

XX AAC19445;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 23520.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX Claim 1; SEQ ID 23520; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or poly(A)-RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.

CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.

SQ Sequence 92 BP; 12 A; 9 C; 15 G; 47 T; 9 other;

Query Match  
 Best Local Similarity 0.5%; Score 26.6; DB 21; Length 92;  
 Matches 36; Conservative 6; Mismatches 21; Indels 0; Gaps 0;

QY 143 aaaaatgaaatgacccaatcgccctccacataaagacactggaacagtggaatg 202  
 Db 72 AAAAAAACMAAAACMAAAACMAAAACMAAAACMAAAACMAAAACMAAAACMAAAATTC 13  
 QY 203 tgg 205  
 Db 12 TGG 10

# RESULT 3

AA086400/c  
 ID AA086400 standard; DNA; 100 BP.

XX AA086400;

DT 16-NOV-1995 (first entry)

DE Human insulin derivative Ala21 A-chain PCR primer #128.

KW Human insulin derivative; A-chain; diabetes; zinc ion complex  
 KW Ala21 PCR primer #128; ss.

OS Synthetic.

PN W09507931-A.

PD 23-MAR-1995.

PF 16-SEP-1994; 94WO-DK00347.

PR 17-SEP-1993; 93DK-0001044.

PR 02-FEB-1994; 94US-0190829.

XX (NOVO) NOVO-NORDISK AS.

PI Andersen AS, Halstrom JB, Havelund S, Jonassen I;

DR Markussen J;

XX WPI; 1995-131314/17.

PT Acylated insulin deriv. which may be present as a zinc ion  
 PT complex - is used to treat diabetes and is rapid acting.

XX Example 1; Page 30; 100pp; English.

CC AA086399 and AA086400 are a pair of primers for the PCR amplification  
 CC of the Asp3 B-chain/Ala21 A-chain human insulin derivative DNA.  
 CC The two subunits of the derivative are covalently connected via  
 CC disulphide bonds between Cys residues A7/B7 and A20/B19. The  
 CC derivative, which may be present as a zinc ion complex, can  
 CC be used as a fast action treatment for diabetes.

SQ Sequence 100 BP; 29 A; 22 C; 23 G; 26 T; 0 other;

Query Match  
 Best Local Similarity 0.5%; Score 26; DB 16; Length 100;  
 Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 4469 aagaatgaagcctgtagaataatgagcttggaatctcttaactactgatatgaa 4528  
 Db 97 AAGTGTACACACGCTAAGGGATGCTTGAACAATGTTGACTTCACTGCTTGTGAC 38



FF 21-APR-1995; 950S-0427097.  
FX  
PR 21-APR-1995; 950S-0427097.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Meagher RB, Summers AO;  
XX WPI; 1997-470112/43.  
DR  
XX DNA encoding mercuric ion reductase in plant-expressible form -- for  
PT producing transgenic plants resistant to toxic heavy metals  
XX  
XX Example 1; Column 43-44; 52pp; English.

CC The present sequence represents a PCR mutagenesis primer used to  
CC mutagenise the mera coding sequence derived from transposon Tn21, to  
CC adapt it for plant expressibility. The nucleic acid molecule produced  
CC is useful for producing transgenic plants that are resistant to toxic  
CC heavy metals (especially mercury) and so can be used remediate and/or  
CC revegetate contaminated areas. The bacterial mera gene is derived from  
CC the transposon Tn21, which was originally isolated from the  
CC incompatibility Group IncFII resistance plasmid NRI.  
XX  
SQ Sequence 99 BP; 29 A; 21 C; 25 G; 24 T; 0 other;

	Query Match	Best Local Similarity	Score 25.2;	DB 18;	Length 99;
	Matches 51;	Conservative	0;	Mismatches 43;	Indels 0;
					Gaps 0;
QY 3026	atgcctcgcataatcttgtaagatgtgtaaggtacaaagccacagctggcgacccaagat				3085
Db 2	aagaccacagatagatgaagcgtgtactgcgtgcattcgcattgaaagcattgaagtc				61
QY 3086	ttagattctacaagaatctgaacctgcattat				3119
Db 62	gtgagcattactcaagccaagcgaagttgcctaat				95

RESULT	8
AAZ28594	
ID	AAZ28594 standard; DNA; 99 BP.
XX	
AC	AAZ28594;
XX	
DT	24-DEC-1999 (first entry)
XX	
DE	Primer 307-339S for Tn21 mera gene.
XX	
KW	Organometal; resistance; transcription; translation; regulation; mera;
KM	transposon; Tn21; merB; mercuroic ion reductase; organomercury lyase; ss;
KW	transgenic plant; detoxification; metal compound; soil; sediment; primer
OS	aquatic environment; bioremediation; contamination; PCR; amplification.
XX	
OS	Synthetic.
OS	Transposon Tn21.
PN	
US	US5965796-A.
PD	12-OCT-1999.
PF	19-JUN-1997; 97US-0878957.
PR	21-APR-1995; 95US-0427097.
PA	(UYGE-) UNIV GEORGIA RES FOUND INC.
PI	Rugh CL, Meagher RB, Summers AO;
DR	WPI; 1999-579950/49.
PT	Metal resistance sequences for producing transgenic plants -



sequence // BP; 27 A; 14 C; 15 G; 21 T; 0 other;

virus (RSV) G gene insert is used with the forward PCR primer (AAS066292) in the construction of bovine/humanPIV3-RSV chimeric viruses. The sequence is described in an invention relating to novel infectious chimeric parainfluenza viruses (PIV3s). The virus comprises a nucleocapsid protein (N), a nucleocapsid phosphoprotein (P), a large genome, or antigenome combined with one or more heterologous gene(s) or more heterologous pathogen(s) encoding one or more antigenic determinant(s) of one or more heterologous pathogen(s) to form a chimeric genome or antigenome. The chimeric PIV is useful in an attenuated vaccine to elicit an immune response against one or more virus(es) selected from human PIV1 (HPV1), HPV2 and HPV3. The chimeric PIV may also elicit a humoral specific immune response against PIV3, measles or RSV. An immunospecific immune may also contain two chimeric PIVs, where the first chimeric PIV elicits an immune response against HPV3 and the second chimeric PIV elicits an immune response against HPV1 or HPV2, and where both the first and second chimeric PIVs elicit an immune response against the non-PIV pathogen. Chimeric HPV3, HPV1 and HPV2 are useful as vaccines to prevent measles and upper or lower respiratory tract infections.



```

RESULT 15
ID AAT65208/c
XX AAT65208 standard; DNA: 71 BP.
AC AAT65208:
XX
XX
DT 10-SEP-1997 (first entry)
XX
DE Transforming growth factor beta-1 binding ligand D 11.
XX
KW Transforming growth factor; beta-1; TGF-beta-1; binding ligand;
KW identification; SELEX; anti-mitogenic; inhibition; cell;
KW Systematic Evolution of Ligands by EXponential enrichment;
KW epithelial; proliferation; diagnosis; treatment; fibroids;
KW kidney; lung; liver; dermal scarring; restenosis; ss.
XX
OS Synthetic.
XX
PN WO9638579-A1.
XX
PD 05-DEC-1996.
XX
PF 30-MAY-1996; 96WO-US08014.
XX
PR 20-MAR-1996; 96US-0618693.
PR 02-JUN-1995; 95US-0458423.
PR 02-JUN-1995; 95US-0458424.
PR 05-JUN-1995; 95US-0465591.
PR 05-JUN-1995; 95US-0465594.
PR 07-JUN-1995; 95US-0479725.
PR 07-JUN-1995; 95US-0479783.
XX
XX (NEXS-) NEXSTAR PHARM INC.
XX
PI Gold L, Janjic N, Pagratlis N, Ringquist S, Toothman PJ;
XX WPI; 1997-034387/03.
XX
XX Identification of nucleic acid ligands to TGF-beta, PDGF and hKGF
XX using SELEX, used in the diagnosis and treatment of proliferative
XX disorders
XX
XX Claim 15; Page 123; 209pp; English.
XX
XX The present sequence, a transforming growth factor beta-1
XX (TGF-beta-1) binding ligand, was identified by Systematic Evolution
XX of Ligands by Exponential enrichment (SELEX). Briefly a candidate
XX mixture of nucleic acids was contacted with TGF-beta-1, and nucleic
XX acids having an increased affinity to TGF-beta-1 partitioned from
XX the remainder of the mixture. The partitioned nucleic acids were
XX then amplified to yield a mixture of nucleic acids enriched for
XX sequences with higher affinity and specifically for binding to
XX TGF-beta-1. The ligand is anti-mitogenic and may be used to inhibit
XX epithelial cell proliferation, or in the diagnosis and treatment of
XX TGF-beta-1 mediated pathological conditions, e.g. fibrotic
XX conditions such as fibroids of the kidney, lung and liver and more
XX acute conditions such as dermal scarring and restenosis.
XX
XX Sequence 71 BP: 15 A; 16 C; 22 G; 18 T; 0 other;
XX
Query Match 0.5%; Score 24.6; DB 18; Length 71;
Best Local Similarity 65.5%; Pred. No. 5.7e+03;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
OY 3958 gaacatgtattagctgattcaagaagcagatcaccatgcatcgaactctcc 4012
DB 56 GAAGACGTAAGTAAGTAATTAATAACATACCCCTGGCGCATGCTCTCC 2

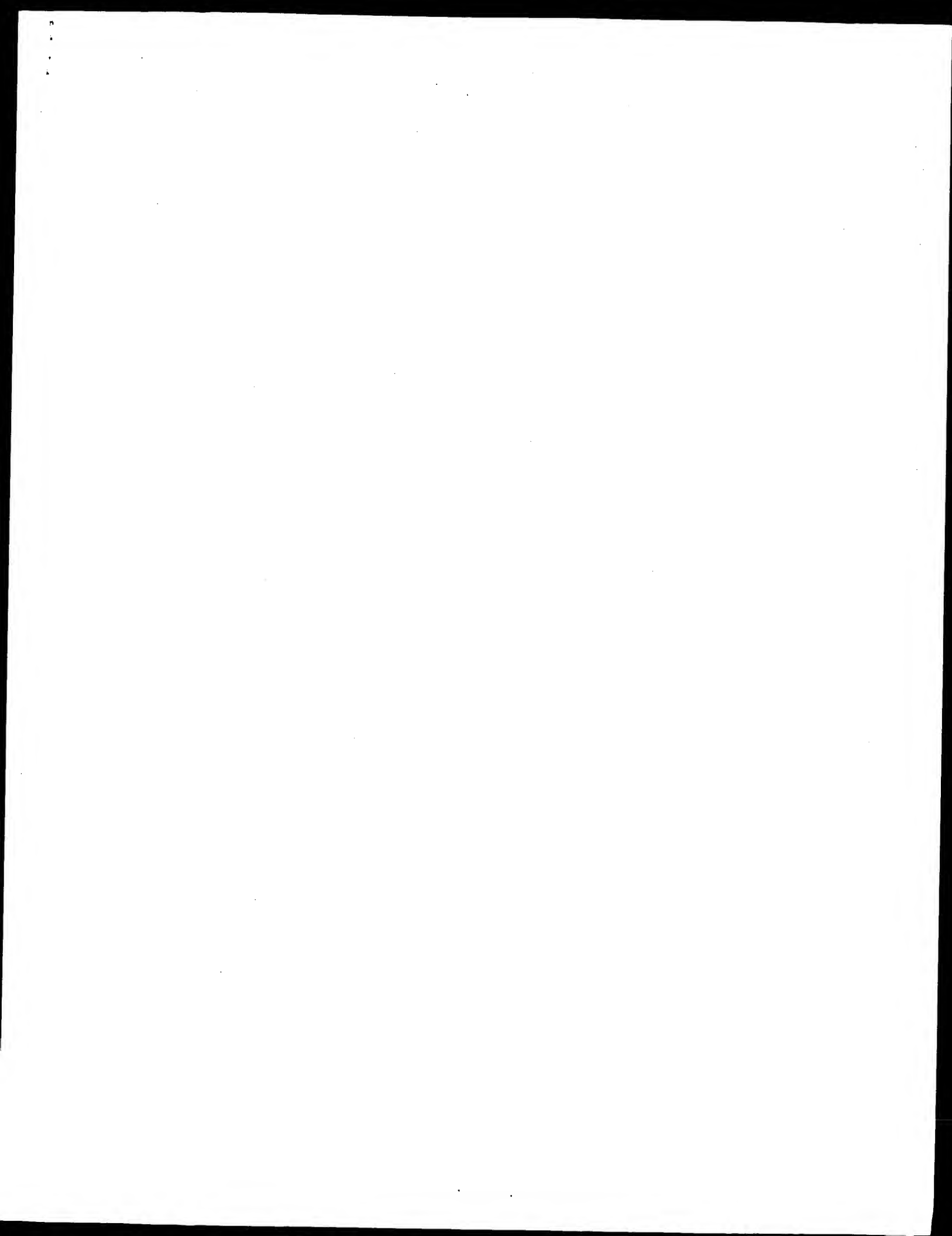
```



Wed May 29 16:25:19 2002

us-09-676-436-3.sz1ml100.rng

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 09:42:16 ; Search time 5865.16 Seconds  
(without alignments)  
17804.019 Million cell updates/sec

Title: US-09-676-436-3  
Sequence: 1 ctgaagaactctcccgatga.....ggttatattagaagaagctc 4990

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 843946

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_da: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sy: \*  
12: gb\_un: \*  
13: gb\_vl: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_mu: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_inv: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

C	1	29.6	0.6	87	3	DDIDDKD	MS9747 D. discoideum
	2	26	0.5	100	6	AR007403	AR007403 Sequence
	3	25.4	0.5	96	5	AF025984	AF025984 Paratrich
	4	25.2	0.5	70	1	AX033167	AX033167 Sequence
	5	25.2	0.5	70	6	AX033159	AX033159 Sequence
	6	25.2	0.5	99	6	165769	165769 Sequence 5
	7	25	0.5	83	4	RABPFXM18	M14473 Rabbit musc
	8	24.8	0.5	92	6	AX173377	AX173377 Sequence
	9	24.6	0.5	71	6	AR140801	AR140801 Sequence
	10	24.6	0.5	71	6	193480	193480 Sequence 78
	11	24.4	0.5	73	6	193107	193107 Sequence 78
	12	24.4	0.5	73	6	S76509	S76509 D18 (A) {h
	13	24.4	0.5	98	3	DROTRANSIN	120950 Drosophila
	14	24.4	0.5	100	6	AR007408	AR007408 Sequence
	15	24	0.5	51	10	U92173	U92173 Mus musculi
	16	24	0.5	66	6	AX207310	AX207310 Sequence
	17	24	0.5	88	9	AB04840511	AB048411 Homo sapi
	18	24	0.5	97	6	AB014621	AB014621 Sequence
	19	24	0.5	97	6	BD010413	BD010413 ChimERIC
	20	24	0.5	97	6	I26754	I26754 Sequence 54
	21	23.8	0.5	82	4	AF294354S5	AF294358 Bos tauru
	22	23.8	0.5	87	14	AF050514	AF050514 Human end
	23	23.8	0.5	90	3	TRBANTATC2	M28495 Trypanosoma
	24	23.8	0.5	91	10	MMU130784	A1130784 Mus muscu
	25	23.8	0.5	94	6	AX326634	AX326634 Sequence
	26	23.8	0.5	94	6	AX326671	AX326671 Sequence
	27	23.8	0.5	94	9	S52152S19	S52191 CD11b-leuko
	28	23.8	0.5	96	5	AF025985	AF025985 Paratrich
	29	23.6	0.5	78	9	HSB305430	AB0305430 Homo sapi
	30	23.6	0.5	87	6	AR062856	AR062856 Sequence
	31	23.6	0.5	88	11	HUMU7770A	L31186 Human SYS U
	32	23.6	0.5	98	6	AX326603	AX326603 Sequence
	33	23.6	0.5	99	6	AX059602	AX059602 Sequence
	34	23.6	0.5	100	5	AF025983	AF025983 Paratrich
	35	23.4	0.5	77	6	A20465	A20465 Oligonucleo
	36	23.4	0.5	79	10	RATPLYC	M32437 Rat/polyoma
	37	23.4	0.5	83	6	AX240922	AX240922 Sequence
	38	23.4	0.5	88	9	S72771	S72771 Immunoglobul
	39	23.4	0.5	89	6	AX015186	AX015186 Sequence
	40	23.4	0.5	99	11	HSPE49C03	AL033823 H. sapiens
	41	23.2	0.5	87	5	CHRM11AS	M19461 Chicken c-m
	42	23.2	0.5	74	5	OCU72578	U72578 Oryctolagus
	43	23.2	0.5	100	5	AF174523	AF174523 Bufo dana
	44	23	0.5	91	9	HSU2336	U32336 Human pre-B
	45	23	0.5	94	6	ARI70294	ARI70294 Sequence

## ALIGNMENTS

RESULT	1	87 bp	DNA	linear	INV 27-APR-1993
DDIDDKD	DDIDDKD				
LOCUS	D. discoideum protein kinase 4 gene, partial cds.				
DEFINITION					
ACCESSION	M59747.1 GI:167723				
VERSION					
KEYWORDS	protein kinase 4.				
SOURCE	Dictyostelium discoideum (strain AX-3) DNA.				
ORGANISM	Dictyostelium discoideum				
REFERENCE	Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.				
AUTHORS	1 (bases 1 to 87)				
TITLE	Haribabu, B. and Dotin, R.P.				
JOURNAL	Identification of a protein kinase multigene family of				
MEDLINE	Dictyostelium discoideum: Molecular cloning and expression of a				
FEATURES	cDNA encoding a developmentally regulated protein kinase				
SOURCE	Proc. Natl. Acad. Sci. U.S.A. 88, 1115-1119 (1991)				
	51142122	location/Qualifiers			
	1..87				
	/organism="Dictyostelium discoideum"				
	/strain="AX-3"				
	/db_xref="taxon:44689"				
	<1..>87				



LOCUS AX033159 70 bp DNA linear PAT 21-SEP-2000  
 DEFINITION Sequence 3 from Patent WO0045176.  
 ACCESSION AX033159  
 VERSION AX033159.1 GI:10280023  
 KEYWORDS  
 SOURCE Escherichia coli.  
 ORGANISM Escherichia coli.  
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 70)  
 Gallusser, A., Karl, J., Lill, H., Stahl, P., Krueger, K. and Borge, A.  
 Method of identifying n-terminal propeptide  
 Patent: WO 0045176-A 3 03-AUG-2000.  
 JOURNAL GALUSSER ANDREAS (DE); KARL JOHANN (DE); LILL HELMUT (DE); STAHL PETER (DE); KRUEGER KERSTIN (DE); BORGIA ANNELEISE (DE); ROCHE DIAGNOSTICS GMBH (DE)  
 FEATURES  
 source location/Qualifiers  
 1..70  
 /organism="Escherichia coli"  
 /db\_xref="taxon:562"

BASE COUNT 10 a 19 c 21 g 20 t

ORIGIN

Query Match 0.5%; Score 25.2; DB 6; Length 70;  
 Best Local Similarity 62.9%; Pred. No. 2.7e+05;  
 Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 1778 agggcgctctgtagaagcagctacacgctgctgtagagagagat 1837  
 |||||  
 DB 8 AGGAGGCTGTCTCAACCTGCACATTCGACAGATTACCTTCAGAGTTCGCTTC 67  
 |||||

OY 1838 tg 1839  
 ||

DB 68 TG 69

RESULT 6  
 LOCUS 165769 99 bp DNA linear PAT 07-OCT-1997  
 DEFINITION Sequence 5 from patent US 5668294.  
 ACCESSION 165769  
 VERSION 165769.1 GI:2482339  
 KEYWORDS  
 SOURCE unknown.  
 ORGANISM unknown.  
 REFERENCE 1 (bases 1 to 99)  
 Mesgher, R.B. and Summers, A.O.  
 Metal resistance sequences and transgenic plants  
 JOURNAL Patent: US 5668294-A 5 16-SEP-1997;  
 FEATURES  
 source location/Qualifiers  
 1..99  
 /organism="unknown"

BASE COUNT 29 a 21 c 25 g 24 t

ORIGIN

Query Match 0.5%; Score 25.2; DB 6; Length 99;  
 Best Local Similarity 54.3%; Pred. No. 2.8e+05;  
 Matches 51; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 3026 atgtccctgaataatgttagagtagagtagacaagccaggtggcgaactcagat 3085  
 |||||  
 DB 2 AAGACCCAGCTATAGCTGAAGCTGTACTGCTATTCGATGGAAGCATGMACTGC 61  
 |||||

OY 3086 ttgatttcaagaagcaattgaacctcttat 3119  
 |||||  
 DB 62 GTGAGCATACTCAAGCAAGCAAGTGGCTATAT 95  
 |||||

RESULT 7  
 RABPFKM18

LOCUS RABPFKM18 83 bp DNA linear NAM 27-APR-1993  
 DEFINITION Rabbit muscle phosphofructokinase gene, exon 18.  
 ACCESSION M14473 J02702  
 VERSION M14473.1 GI:165622  
 KEYWORDS phosphofructokinase.  
 SEGMENT 18 OF 22  
 SOURCE Rabbit (New Zealand) DNA, clone lambda-Charon 4APK.  
 ORGANISM Oryctolagus cuniculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE 1 (bases 1 to 83)  
 Lee, C.-P., Kao, M.-C., French, B.A., Putney, S.D. and Chang, S.H.  
 The rabbit muscle phosphofructokinase gene: Implications for protein structure, function and tissue specificity  
 JOURNAL J. Biol. Chem. 262, 4195-4199 (1987)  
 MEDLINE 87166033  
 COMMENT Draft entry and computer-readable sequence for [1] kindly provided by S.H.Chang, 02-FEB-1987.  
 FEATURES  
 source location/Qualifiers  
 1..83  
 /organism="Oryctolagus cuniculus"  
 /db\_xref="taxon:9986"  
 <1..15  
 /note="PK Intron Q"  
 16..77  
 /number=18  
 78..>83  
 /note="PK Intron R"  
 23 a 27 g 13 c 20 t  
 About 719 bp after segment 17.

BASE COUNT 23 a 27 g 13 c 20 t

ORIGIN

Query Match 0.5%; Score 25; DB 4; Length 83;  
 Best Local Similarity 61.5%; Pred. No. 3.1e+05;  
 Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 1737 gagtaataagcagctgtagagagtagagagtagagagtagagagtagagag 1796  
 |||||  
 DB 18 GATGCTGACACCTGTGTCGCAAGATGACGACGCTGTAAGAGCGCTGTGCTGAG 77  
 |||||

OY 1797 gcaat 1801  
 |||||

DB 78 GTACT 82

RESULT 8  
 LOCUS AX173377 92 bp DNA linear PAT 03-JUL-2001  
 DEFINITION Sequence 31 from Patent WO0142445.  
 ACCESSION AX173377  
 VERSION AX173377.1 GI:14598152  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.  
 REFERENCE 1 (bases 1 to 92)  
 Murphy, B.R., Collins, P.L., Schmidt, A.C., Durbin, A.P., Skladopoulos, M.H. and Tao, T.  
 Use of recombinant parainfluenza viruses (pIv) as vectors to protect against infection and disease caused by pIv and other human pathogens  
 Patent: WO 0142445-A 31 14-JUN-2001;  
 JOURNAL The Secretary of the Department of Health and Human Services (US)  
 FEATURES  
 source location/Qualifiers  
 1..92  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="Reverse primer for RSV A G gene insert"

BASE COUNT 20 a 12 c 26 g 34 t

ORIGIN

Query Match 0.5%; Score 24.8; DB 6; Length 92;

Best Local Similarity 63.3%; Pred. No. 3.5e+05;  
Matches 38; Conservation 0.0011; W. 1.0000

[illegible]

D<sub>b</sub> 80 TTCTCTCCACCCAAACACACACAGCAGTAGTAAAAAGTTAAGAAAACTTAGATTAA 21

	RESULT	9		
	ARI40801/c			
	LOCUS			
	DEFINITION	71 bp	DNA	
	Sequence	78 from patent US 6207816.		
	ACCESSION	ARI40801		linear
	VERSION	ARI40801.1		PAT 16-JUN-2001
	KEYWORDS	GI:14483297		.
	SOURCE	. unknown.		
	ORGANISM	unknown.		

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1 (bases 1 to 71)	Gold, L., Janjic, N. and Pagratis, N.	High affinity oligonucleotide ligands to growth factors	Patent: US 6207816-A 78 27-MAR-2001,	Location/Qualifiers	1 71

BASE COUNT	/organism="unknown"
15 a	16 c
22 g	18 t

Query Match	0.58;	Score	24.6;	DB	6;	Length	71;
Best Local Similarity	65.5%;	Best	no. 3	0.005			

Matches	36;
Conservative	0:
Mismatches	19;
Freq. NO.	3.9e+05;

Oy 3958 gaacatgattagctgtatccaagaatccattcgatacaagtcttc 4012  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
56 GAAGACGTAGTAGTACTATTAAACAATACCccccctGCGCATCTCCTCC 2  
Db

RESULT 10			
193480/c			
LOCUS	193480	71 bp	DNA
DEFINITION	Sequence 78 from patent US 5731144.		
ACCESSION	U03490		
FEATURES			linear
			PAT 01-DEC-1998

REFERENCE 1 (Pages 1 to 71)  
AUTHORS Toothman,P.J., Ringquist,S. and Gold,L.  
TITLE High affinity TGF beta<sub>2</sub> nucleic acid ligands  
JOURNAL Patent: US 5731144-A 78 24-MAR-1998;  
FEATURES location/Qualifiers

source	1. .71	
BASE COUNT	15 a	16 c
ORIGIN		22 g
		18 t

Query Match	0.5%	Score 24.6;	DB 6;	Length 71;
Best Local Similarity	65.5%	Pred. No. 3.9e+05;		
Matches	36;	Conservative	0;	Mismatches 10;

**QY** 3958 gaacatgtattagcgcgtatccaagcagatcaccaatgcatcaactctcc 4012  
||| || | | |||| || | | | | | ||  
**Db** 56 GAAGACGTAACTAGTATTAATAAACATACCCTCCCTGGCGCATGCTCCTCC 2

RESULT	11		
195107/c			
LOCUS			
195107	71 bp	DNA	
Sequence	78	from patent US 573124.	
DEFINITION			
195107			
			linear
			PAT 01-DEC-1998

VERSION	I95107.1	GI:3939577
KEYWORDS	.	
SOURCE	unknown.	

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source
1	(bases 1 to 71)	unclassified.			
	Toothman, P.J., Ringquist, S. and Gold, L.	High affinity TRF-beta nucleic acid ligands and inhibitors	Patent: US 5731424-A 78 24-MAR-1998;		
		Location/Qualifiers			1. .71

BASE COUNT	15 a	16 c	22 g	18 t
ORIGIN	/organism="unknown"			

Query Match	0.5%;	Score 24.6;	DB 6;	Length 71;
Best Local Similarity	65.5%;	Pred. No. 3.9e+05;		
Matches 36;	Conservative 0;	Mismatches 19;	Totals 0;	Coverage 0

Qy 3958 gaacatgattagcgtgatccaagaataccattcgacaacgctctccc 4012  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 56 GAAGACGTAACTAATAATTAAAAATATCCCCCCTGCGCATCGTCGCC 2

RESULT 12  
S76509/c

LOCUS	73 bp	DNA	linear	PRI 07-MAY-1993
DEFINITION	DIS8 (A) [hypervariable minisatellite MS32] [human, individual MACH, Genomic, 73 nt].			

```

VERSION      S76509.1  GI:243382
KEYWORDS
SOURCE      human individual MA

```

REFERENCE  
1 (bases 1 to 73)  
Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.  
Chordata; Vertebrata; Euteleostomi;  
Cnidaria; Cephalopoda; Mollusca; Annelida; Arthropoda; Insecta; Metazoa;

**AUTHORS** Monckton, D.G. and Jeffreys, A.J.  
**TITLE** Minisatellite 'isolate' discrimination in pseudohomozygotes by single molecule PCR and variant repeat mapping  
**JOURNAL** Genomics 11 (2), 465-467 (1991)  
**PAGE** 92120671  
**KEYWORDS**

**REMARK** GenBank staff at the National Library of Medicine created this entry [NCBI gi|537509] from the original journal article. This sequence comes from Figure 2.

EXPLANES	Location/qualifiers
source	1. .73

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
1..73
gene

```

BASE COUNT	41 a	0 c	0 g	32 t
IRIGIN				

Query Match	0.5%	Score 24.4	DB 9	length 73
Best Local Similarity	60.6%	Pred. No. 4.5e+05		
Matches 40; Conservative	0	Mismatches 26	Indels 0	Gaps 0

[illegible]

4	49/4	ctatt
7	12	TTATT

RESULT	LOCUS	98 bp	DNA	linear	INV 05-JAN-1994
13	DROTRANSIN/c				
	DROTRANSIN				

```

DEFINITION   Drosophila melanogaster DNA sequence, complete insertion sequence.
ACCESSION    I20950
VERSION      I20950.1 GI:304832
KEYWORDS     insertion sequence
SOURCE       Drosophila melanogaster (individual isolate B20.3) DNA.
ORGANISM     Drosophila melanogaster
REFERENCE    1 (bases 1 to 98)
AUTHORS      Engels, W.R. and Johnson-Schultz, D.M.
TITLE        Deletion white allele. Transposase-induced derivative of
             white-hd80k17
JOURNAL      Unpublished (1993)
REFERENCE    2 (bases 1 to 98)
AUTHORS      Johnson-Schultz, D.M. and Engels, W.R.
TITLE        P-element-induced interallelic gene conversion of insertions and
             deletions in Drosophila melanogaster
JOURNAL      Mol. Cell. Biol. 13, 7006-7018 (1993)
MEDLINE      94019372
FEATURES     Location/Qualifiers
             source          1..98
                           /organism="Drosophila melanogaster"
                           /isolate="B20.3"
                           /db_xref="taxon:7227"
BASE COUNT   37 a      2 c      1 g      58 t
ORIGIN
Query Match 0.5%; Score 24.4; DB 3; Length 98;
Best Local Similarity 56.1%; Pred. No. 4.5e+05;
Matches 46; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1362 tctactctatctatagaccattgtagacaaagcactgaagcagatggggttagaana 1421
Db 85 TATATAATATATATATATATATATATATATATATATATATATATATATATATATATAA 26
QY 1422 gtaatttaagacttacaag 1443
Db 25 TTATATTTAATATATATATATCATG 4
RESULT 14
LOCUS        AR007408 100 bp DNA linear PAT 04-DEC-1998
DEFINITION   Sequence 9 from patent US 5750497.
ACCESSION    AR007408
VERSION      AR007408.1 GI:3966892
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 100)
AUTHORS      Havelund, S., Halstrøm, J., Jonassen, I., Andersen, A., Sloth,
             and Markussen, J.
TITLE        Acylated insulin
JOURNAL      Patent: US 5750497-A 9 12-MAY-1998;
FEATURES     Location/Qualifiers
             source          1..100
                           /organism="unknown"
BASE COUNT   29 a      23 c      22 g      26 t
ORIGIN
Query Match 0.5%; Score 24.4; DB 6; Length 100;
Best Local Similarity 54.4%; Pred. No. 4.5e+05;
Matches 49; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 4469 aagaatgaagccttagaataagacttggaataatctcttaactactatgtatgtaa 4528
Db 97 AAGTCGTACGACGCTGATGCGTGAACAAATGTTGACTTCCTATCTCTTTGTAG 38
QY 4529 tattacataaagacttgctgagaagcag 4558

```

```

Db 37 CAATTGAAAACACTGTGTAGACGACG 8
RESULT 15
LOCUS        U92173 51 bp mRNA linear ROD 19-FEB-1998
DEFINITION   Mus musculus clone 1C3 T cell receptor beta chain mRNA, partial
KEYWORDS     cds
ACCESSION    U92173
VERSION      U92173.1 GI:2894946
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    1 (bases 1 to 51)
AUTHORS      Eukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        T cell receptor diversity in CTLs specific for the CTT-1 and CTT-2
             minor histocompatibility antigens
JOURNAL      J. Immunol. 159 (6), 2606-2615 (1997)
MEDLINE      97444147
REFERENCE    2 (bases 1 to 51)
AUTHORS      Johnston, S.L. and Wettstein, P.J.
TITLE        Direct Submission
JOURNAL      Submitted (06-MAR-1997) Immunology, Mayo Clinic, 200 1st St S.W.,
             Rochester, MN 55905, USA
FEATURES     Location/Qualifiers
             source          1..51
                           /organism="Mus musculus"
                           /strain="C57BL/6"
                           /db_xref="taxon:10090"
                           /clone="1C3"
                           /cell_type="cytotoxic T lymphocyte"
                           /note="minor CTT-1 antigen specific CDR3 region"
                           /product="T cell receptor beta chain"
                           /protein_id="AAC02878.1"
                           /db_xref="GI:2894947"
                           /translation="YCTCSDPDVSGNTLYF"
BASE COUNT   12 a      13 c      11 g      15 t
ORIGIN
Query Match 0.5%; Score 24; DB 10; Length 51;
Best Local Similarity 75.0%; Pred. No. 5.6e+05;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1513 gaattccagatcccatgtggttcagattatgtcagt 1552
Db 41 GTATTTCCAGATVACCTGTGGGATCATCTGCAGGTGCACT 2

```

Search completed: May 29, 2002, 12:55:02  
 Job time: 11566 sec

Wed May 29 16:25:19 2002

us-09-676-436-3.sz1ml100.rge

Page 6